Nonparametric Frontier Analysis using Stata

Oleg Badunenko
University of Cologne
Cologne/Germany
obadunen@uni-koeln.de

Pavlo Mozharovskyi
Agrocampus Ouest
Rennes/France
pavlo.mozharovskyi@univ-rennes1.fr

Abstract. This article describes five new Stata commands that estimate and provide statistical inference in nonparametric frontier models. First two commands, `tenonradial` and `teradial`, estimate data envelopment models where nonradial and radial technical efficiency measures are computed (Färe 1988; Färe and Lovell 1994; Färe et al. 1994a). Technical efficiency measures are obtained by solving linear programming problems. The rest of the commands, `teradialbc`, `nptestind`, and `nptestrts`, give tools for making statistical inference regarding radial technical efficiency measures (Simar and Wilson 1998, 2000, 2002). The article provides brief overview of the nonparametric efficiency measurement, as well as the description of syntax and options of new commands. Additionally, an example showing the capabilities of new commands is provided. Finally, a small empirical study of productivity growth is performed.

Keywords: st0001, `tenonradial`, `teradial`, `teradialbc`, `nptestind`, `nptestrts`, Nonparametric efficiency analysis, Data Envelopment Analysis, Technical efficiency, Radial measure, Nonradial measure, Linear programming, Bootstrap, Subsampling bootstrap, Smoothed bootstrap, Bias-correction

1 Introduction

The concept of efficiency is at the core of production economics. Beginning with the pioneering work by Cobb and Douglas (1928), there were many attempts to parameterize the production process: e.g., Leontieff, Constant Elasticity of Substitution, transcendental logarithmic production and cost functions. Conceptually, however, researchers looked at the "average" input-output relationship assuming no inefficiency. Yet, it was no longer plausible to assume that all units are homogeneous, that is, operating at the same level of efficiency. Among the first to offer an appropriate modification was Farrell (1957), who built up on the concept of efficiency postulated by Koopmans (1951) and Debreu (1951) and put forward a foundation, which has become a distinct field in economics—the efficiency analysis. Färe (1988); Färe et al. (1994a), Färe and Primont (1995) provide many insights into nonparametric efficiency measurement.

Data Envelopment Analysis (DEA), a leading analytical technique for measuring relative efficiency, has been widely used by both academic researchers and practitioners in evaluating the efficiency of decision making units in terms of converting inputs into outputs. Researchers choose this technique because it does not impose *a priori* functional form and allows for multiple output technologies.

Although the DEA method is typically considered to be deterministic, the efficiency is still computed relatively to estimated and not true frontier. The efficiency scores obtained from a finite sample are subject to sampling variation of the estimated frontier. Simar and Wilson (1998, 2000, 2002) have laid out a statistical model and proposed consistent bootstrap procedures to provide statistical inference regarding technical efficiency measures in nonparametric frontier models.

The estimation of DEA model can be readily performed in Stata using user-written command `dea` (Ji and Lee 2010). However, `dea` is limited in its capability and is slow with even moderate data sets. We provide time comparison of `dea` and our command. The five new Stata commands described here estimate and provide statistical inference in nonparametric frontier models. `tenonradial` and `teradial` estimate data envelopment models where nonradial and radial technical efficiency measures are computed (Färe 1988; Färe and Lovell 1994; Färe et al. 1994a). `teradialbc`, `nptestind`, and `nptestrts` allow making statistical inference regarding radial technical efficiency measures (Simar and Wilson 1998, 2000, 2002).

The remainder of the paper is structured as follows: section 2 provides an overview of nonparametric frontier models; sections 3–7 contain the syntax and explain the options of new commands; section 8 illustrates the capabilities of new commands using data set for Program Follow Through at 70 US Primary Schools and performs the analysis of the changes in productivity for 52 countries using Penn World Tables; section 9 discusses the features and limitations of new commands; section 10 emphasizes the difference between our and `dea` command, and section 11 concludes.

2 Nonparametric frontier analysis

This section introduces two types of nonparametric efficiency measurement, radial and nonradial. Further, recent statistical developments regarding radial measure are discussed. The exposition here has only an introductory nature. For more details refer to the cited papers and books.

2.1 Radial efficiency analysis

Our measures of technical efficiency for the production data points are conventional radial Debreu-Farrell measure of efficiency loss (Debreu 1951; Farrell 1957). For each data point k $(k = 1, \ldots, K)$ vector $x_k = (x_{k1}, \ldots, x_{kN}) \in \Re^N$ denotes N inputs, vector $y_k = (y_{k1}, \ldots, y_{kM}) \in \Re^M$ denotes M outputs. We assume that under technology T the data (y, x) are such that outputs are producible by inputs,

$$T = \{(x, y) : y \text{ are producible by } x\}. \tag{1}$$

The technology is fully characterized by its production possibility set,

$$P(x) \equiv \{y : (x, y) \in T\} \tag{2}$$

or input requirement set,

$$L(y) \equiv \{x : (x, y) \in T\}. \tag{3}$$

Conditions (2) and (3) imply that the available outputs and inputs are feasible. The upper boundary of the production possibility set and lower boundary of the input requirement set define the frontier. How far a given data point is from the frontier represents its efficiency. In output-based radial efficiency measurement, the amount of necessary (proportional) expansion of outputs to move a data point to a boundary of the production possibility set $P(x)$ serves a measure of technical efficiency. In input-based radial efficiency measurement, it is the amount of necessary (proportional) reduction of inputs to move a data point to a boundary of the input requirement set $L(y)$.

Empirically, technical efficiencies are estimated via activity analysis models. These are widely known as Data Envelopment Analysis (DEA). For K data points, M outputs and N inputs an estimate of the radial Debreu-Farrell output-based measure of technical efficiency can be calculated by solving a linear programming problem for each data point k $(k = 1, \ldots, K)$:

$$\begin{aligned}\hat{F}_k^o(y_k, x_k, y, x|\text{CRS}) = \max_{\theta,z} \theta \qquad & \qquad (4)\\ \text{s.t.} \sum_{k=1}^{K} z_k y_{km} &\geq y_{km}\theta_m, m = 1, \cdots, M,\\ \sum_{k=1}^{K} z_k x_{kn} &\leq x_{kn}, n = 1, \cdots, N,\\ z_k &\geq 0.\end{aligned}$$

y is $K \times M$ matrix of available data on outputs, x is $K \times N$ matrix of available data on inputs. The estimate of $P(x)$ is the smallest convex free-disposal hull that envelops the observed data, and upper boundary of which is a piece-wise linear estimate of the true best-practice frontier of $P(x)$. Equation (4) gives us constant returns to scale (CRS) specification. Other returns to scale are modeled by adjusting process operating levels z_k's; for variable returns to scale (VRS) a convexity constraint $\sum_{k=1}^{K} z_k = 1$ is added,[1] while for non-increasing returns to scale (NIRS), $\sum_{k=1}^{K} z_k \leq 1$ inequality is added[2] to set of restrictions of linear programming problem in equation (4).

To facilitate the discussion, figures 1 and 2 present hypothetical one-input one-output production processes with three different technologies CRS, VRS and NIRS. Conceptually, in figure 1 (2) the vertical (horizontal) distance from a data point (x_i, y_i) or (x_j, y_j) to CRS/VRS/NIRS best-practice frontier stands for output-based (input-based) technical efficiency under assumption of CRS/VRS/NIRS technology. In a multi-dimensional case, the required distance is the radial path from a data point that is parallel to axes along which all outputs (inputs) are measured.

1. This equality ensures that data point k is compared only to data points of similar size; under CRS assumption, data points of different sizes might be compared to one another.
2. This inequality ensures that data point k is not compared to other data points that are considerably larger, but maybe compared to smaller data points.

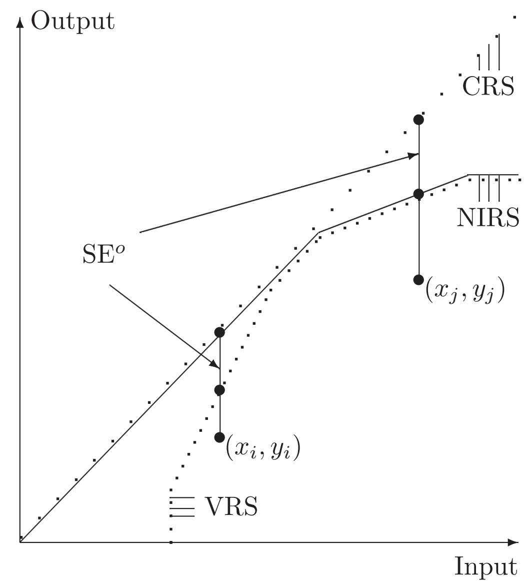

Figure 1: Output-based technical and scale efficiency

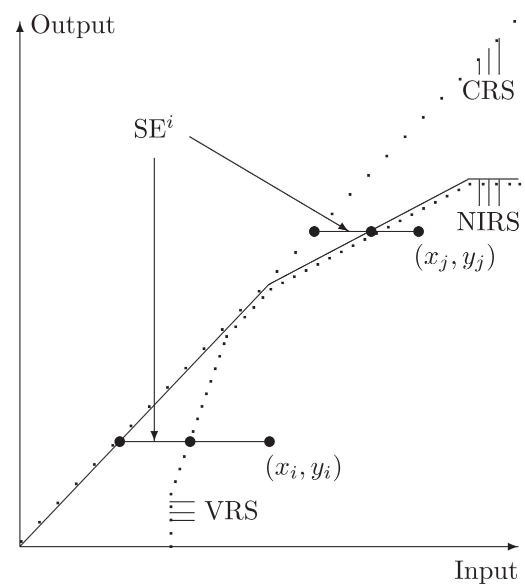

Figure 2: Input-based technical and scale efficiency

2.2 Nonradial efficiency analysis

For data point (y_k, x_k), radial measure expands (shrinks) all M outputs $y_k = (y_{k1}, \ldots, y_{kM})$ (N inputs $x_k = (x_{k1}, \ldots, x_{kN})$) proportionally until the frontier is reached. At the reached frontier point, some but not all outputs (inputs) can there expanded (shrunk) while remaining feasible. If such possibility is available for a given data point k for output m (input n), the reference point $F^o_k(y_k, x_k) \times y_{mk}$ ($F^i_k(y_k, x_k) \times x_{nk}$) is said to have slack in output y_m (input x_n). Nonradial measure of technical efficiency, the Russell measure, accommodates such slacks (Färe and Lovell 1994; Färe et al. 1994a). The output-based nonradial measure for data point$j$ is defined by

$$RM^o_k(y_k, x_k, y, x|\mathrm{CRS}) = \max \left\{ \begin{array}{r} M^{-1} \sum\limits_{m=1}^{M} \theta_m\text{: } (\theta_1 y_{k1}, \ldots \theta_M y_{kM}) \in P(x), \\ \theta_m \geq 0, m = 1, \ldots, M \end{array} \right\}. \tag{5}$$

The input-based counterpart is given by

$$RM^i_k(y_k, x_k, y, x|\mathrm{CRS}) = \min \left\{ \begin{array}{r} N^{-1} \sum\limits_{n=1}^{N} \lambda_n\text{: } (\lambda_1 x_{k1}, \ldots \lambda_N y_{kN}) \in L(y), \\ \lambda_n \geq 0, n = 1, \ldots, N \end{array} \right\}. \tag{6}$$

The output-based Russell measure can be calculated for positive outputs as a solution to the linear programming problem

$$\begin{aligned} \widehat{RM}^o_k(y_k, x_k, y, x|\mathrm{CRS}) = M^{-1} \max_{\theta, z} \sum_{m=1}^{M} \theta_m \qquad \qquad (7) \\ \text{s.t. } \sum_{k=1}^{K} z_k y_{km} \geq y_{km}\theta_m, m = 1, \cdots, M, \\ \sum_{k=1}^{K} z_k x_{kn} \leq x_{kn}, n = 1, \cdots, N, \\ z_k \geq 0. \end{aligned}$$

and input-based Russell measure can be calculated for positive inputs as a solution to the linear programming problem

$$\begin{aligned} \widehat{RM}^i_k(y_k, x_k, y, x|\mathrm{CRS}) = N^{-1} \min_{\theta, z} \sum_{n=1}^{N} \lambda_n \qquad \qquad (8) \\ \text{s.t. } \sum_{k=1}^{K} z_k y_{km} \geq y_{km}, m = 1, \cdots, M, \\ \sum_{k=1}^{K} z_k x_{kn} \leq x_{kn}\lambda_n, n = 1, \cdots, N, \\ z_k \geq 0. \end{aligned}$$

If output $y_{km} = 0$ ($x_{kn} = 0$), the linear programming problem in 7 (8) is modified and θ_m (λ_n) is set to 1.

The Russell measure allows for nonproportional expansions (reductions) in each positive output (input). The nonradial output- (input-) based Russell measure collapses to the radial measure when $\theta_m = \theta, \forall m$, where $y_{km} > 0$ ($\lambda_n = \lambda, \forall n$, where $x_{kn} > 0$). However, since the Russell measure can expand (shrink) an output (input) vector at most (least) as far as the radial measure can, we have the result that

$$1 \geq \hat{F}_k^o(y_k, x_k, y, x|\text{CRS}) \geq \widehat{RM}_k^o(y_k, x_k, y, x|\text{CRS}) \tag{9}$$

and

$$0 < \widehat{RM}_k^i(y_k, x_k, y, x|\text{CRS}) \leq \hat{F}_k^i(y_k, x_k, y, x|\text{CRS}) \leq 1. \tag{10}$$

Technologies under nonincreasing and variable returns to scale can be modeled by imposing respective restrictions on the intensity vector, z, in the piecewise linear technology, that is in (7) and (8). Then the Russell measure can be calculated relative to these technologies.

In case of one input (output), the input (output)-based Russell measure is equal to Debreu-Farrell radial measure of technical efficiency.

2.3 Statistical inference in radial frontier model

Although the DEA method is typically considered to be deterministic, the efficiency is still computed relatively to estimated and not true frontier. The efficiency scores obtained from a finite sample (in equation (4) from K data points) are subject to sampling variation of the estimated frontier. The estimated technical efficiency measures are too optimistic, due to the fact that the DEA estimate of the production set is necessarily a weak subset of the true production set under standard assumptions underlying DEA. The statistical inference regarding the radial DEA estimates can be provided via bootstrap technique. The details of the concept and implementation of the bootstrap mechanism are given in Simar and Wilson (1998, 2000); Kneip et al. (2008). The bootstrapping procedure allows to estimate the bias and the confidence interval of the original estimate. Badunenko et al. (2012) study statistical properties of the bias-corrected estimator in finite samples.

2.4 Type of the bootstrap for statistical inference

The bootstrapping technique mentioned in the previous section relies on several assumptions. In output-based efficiency measurement, the major assumption depends on whether the estimated output-based measures of technical efficiency are independent of the mix of outputs. In input-based efficiency measurement, the major assumption depends on whether the estimated input-based measures of technical efficiency are independent of the mix of inputs. This dependency is testable given the assumption of returns to scale of the global technology (Wilson 2003). If output-based measures of technical efficiency are independent of the mix of outputs, the smoothed homogeneous bootstrap can be used. This type of the bootstrap is not computer intensive. If on the contrary, output-based measures of technical efficiency are *not* independent of the mix of outputs, the heterogenous bootstrap must be used to provide valid statistical inference. The latter type of bootstrap is quite computer demanding and may take a while for large data sets.

2.5 Returns to scale and scale analysis

The assumption regarding the global technology is crucial in DEA. Depending on this assumption equation (4) and resulting measures of technical efficiency will vary. The assumption about returns to scale should be made using prior knowledge about the particular industry. If this knowledge does not suffice, or is not conclusive, the returns to scale assumption can be tested econometrically. Moreover, if technology is not CRS globally, estimating measure of technical efficiency under CRS will lead to inconsistent results (Simar and Wilson 2002).

The measures of radial technical efficiency in equation (4) under CRS, NIRS, and VRS can be used to calculate the measures of scale efficiency, originally proposed by Färe and Grosskopf (1985),

$$S_k^o(y_k, x_k) = \frac{\hat{F}_k^o(y_k, x_k, y, x|\mathrm{CRS})}{\hat{F}_k^o(y_k, x_k, y, x|\mathrm{VRS})}, \tag{11}$$

and

$$S_k^{o*}(y_k, x_k) = \frac{\hat{F}_k^o(y_k, x_k, y, x|\mathrm{NIRS})}{\hat{F}_k^o(y_k, x_k, y, x|\mathrm{VRS})} \tag{12}$$

for output-based analysis and

$$S_k^i(y_k, x_k) = \frac{\hat{F}_k^i(y_k, x_k, y, x|\mathrm{CRS})}{\hat{F}_k^i(y_k, x_k, y, x|\mathrm{VRS})}, \tag{13}$$

and

$$S_k^{i*}(y_k, x_k) = \frac{\hat{F}_k^i(y_k, x_k, y, x|\mathrm{NIRS})}{\hat{F}_k^i(y_k, x_k, y, x|\mathrm{VRS})} \tag{14}$$

for input-based analysis. Scale efficiency, S_k^o measures how close is the data point (y_k, x_k) to potentially optimal scale, also known as maximum productive scale size (MPSS), the portion of the frontier where CRS and VRS frontiers coincide in figures 1 and 2 (denoted by SE^o and SE^i, respectively). If $S_k^o(y_k, x_k) = 1$ ($S_k^i(y_k, x_k) = 1$ in input-based efficiency measurement), a data point (y_k, x_k) is scale efficient. If $S_k^o(y_k, x_k) > 1$ ($S_k^i(y_k, x_k) < 1$ in input-based efficiency measurement), a data point (y_k, x_k) is scale inefficient due to operating under the decreasing returns portion of technology if $S_k^{o*}(y_k, x_k) = 1$ ($S_k^{i*}(y_k, x_k) = 1$ in input-based efficiency measurement) or due to operating under the increasing returns portion of technology if $S_k^{o*}(y_k, x_k) > 1$ ($S_k^{i*}(y_k, x_k) < 1$ in input-based efficiency measurement).

On the one hand, if global technology T in equation (1) represents CRS, the VRS estimator is less efficient than CRS. On the other hand, if global technology T in equation (1) is not CRS at some mix of outputs (inputs), CRS estimator is inconsistent. Therefore Simar and Wilson (2002) suggest the following tests

$$\text{Test \#1:} \quad \begin{aligned} &H_0 : T \text{ is globally CRS} \\ &H_1 : T \text{ is VRS.} \end{aligned}$$

If null hypothesis H_0 is rejected, that is, technology is not CRS everywhere, the following test with less restrictive null hypothesis may be performed

$$\text{Test \#2:} \quad \begin{aligned} &H_0' : T \text{ is globally NIRS} \\ &H_1 : T \text{ is VRS.} \end{aligned}$$

Using scale efficiency measures for all K data points, the statistics for testing Test #1 and Test #2 are defined by

$$\hat{S}_{2n}^{o} = \frac{\sum_{k=1}^{K} \hat{F}_k^o(y_k, x_k, y, x|\text{CRS})}{\sum_{k=1}^{K} \hat{F}_k^o(y_k, x_k, y, x|\text{VRS})} \tag{15}$$

and

$$\hat{S}_{2n}^{o\prime} = \frac{\sum_{k=1}^{K} \hat{F}_k^o(y_k, x_k, y, x|\text{NIRS})}{\sum_{k=1}^{K} \hat{F}_k^o(y_k, x_k, y, x|\text{VRS})}. \tag{16}$$

The idea of testing the null hypothesis that the technology is globally CRS versus the alternative hypothesis that the technology is globally VRS, Test #1, boils down to testing how far the test statistic (15) is from its bootstrap analog. This statistic represents the ratio of the average measures of technical efficiency under assumption of VRS and CRS technologies. If null hypothesis is true, then average distance between VRS and CRS frontiers is small. If alternative hypothesis is true, then distance between VRS and CRS frontiers on average is large—the null hypothesis H_0 is rejected if $\hat{S}_{2n}^{o}$ is significantly larger than 1 ($\hat{S}_{2n}^{i}$, defined similar to (15) smaller than 1 in input-based efficiency measurement). If H_0 is rejected, Test #2 can be performed to test the null hypothesis H_0' that the technology is globally NIRS versus the alternative hypothesis that the technology is globally VRS. Analogously to Test #1, if null hypothesis H_0' is true, then average distance between VRS and NIRS frontiers is small. If alternative hypothesis is true, then distance between VRS and NIRS frontiers on average is large—the null hypothesis H_0' is rejected if $\hat{S}_{2n}^{o\prime}$ is significantly larger than 1 ($\hat{S}_{2n}^{i\prime}$, defined similar to (16) smaller than 1 in input-based efficiency measurement).

Taking into account the importance of returns to scale assumption for DEA estimator, this data-driven test is advised to be performed before applying any DEA model.

Additionally, this testing procedure can be used to perform the scale analysis for each data point. The CRS assumption is only feasible when all data points are operating at an optimal scale: i.e., when scale efficiency is unity. However, for many reasons (e.g., imperfect competition, financial constraints) it is more appropriate to assume variable returns to scale (see Coelli et al. 2002, for history and development of the this stream). Assuming CRS when VRS should be assumed in reality overestimates technical efficiency estimate exactly by scale efficiency. Therefore, performing the individual returns to scale test is fairly important in case of scale efficiency analysis.

First, for each data point k, the null hypothesis of Test $\#1_k$ that measures of technical efficiency are equal under constant and variable returns to scale or $S_k^o(y_k, x_k) = 1$ against an alternative hypothesis that $S_k^o(y_k, x_k) > 1$ ($S_k^i(y_k, x_k) < 1$ in input-based case) is tested.[3] Since by definition $S_k^o(y_k, x_k) \geq 1$ ($S_k^i(y_k, x_k) \geq 1$ in input-based case), such null hypothesis is rejected if $S_k^o(y_k, x_k)$ is significantly greater than 1 ($S_k^i(y_k, x_k) \leq 1$ in input-based case). The data point $S_k^o(y_k, x_k)$, for which this null hypothesis is rejected, $S_k^o(y_k, x_k) > 1$ ($S_k^i(y_k, x_k) < 1$ in input-based case), is said to be scale inefficient. Second, for all scale inefficient data points a null hypothesis of Test $\#2_k$ that the measures of technical efficiency are equal under nonincreasing and variable returns to scale or $S_k^{o*}(y_k, x_k) = 1$ ($S_k^{i*}(y_k, x_k) = 1$ in input-based case) against an alternative hypothesis that $S_k^{o*}(y_k, x_k) > 1$ ($S_k^{i*}(y_k, x_k) < 1$ in input-based case) can be performed. The Test $\#2_k$ concludes that data point (y_k, x_k) is operating under increasing returns to scale (such as a data point (x_i, y_i) in terms of figure 1 or 2) if $S_k^{o*}(y_k, x_k)$ is significantly larger than 1 ($S_k^{i*}(y_k, x_k) < 1$ in input-based case), or is operating under decreasing returns to scale (such as a data point (x_j, y_j) in terms of figure 1) otherwise. All tests in this subsection are tests based on bootstrap techniques mentioned in the previous section.

3 The `tenonradial` command

`tenonradial` uses reduced linear programming to compute the nonradial output- or input-based measure of technical efficiency, which is known as the Russell measure. In input-based nonradial efficiency measurement, this measure allows for non-proportional (different) reductions in each positive input, and this is what permits it to shrink an input vector all the way back to the efficient subset. In output-based nonradial efficiency measurement, the Russell measure allows for non-proportional (different) expansions of each positive output

3. Note, that $S_k^o(y_k, x_k)$ is the test statistic of Test $\#1_k$.

3.1 Syntax

`tenonradial` *outputs* = *inputs* [(*ref_outputs* = *ref_inputs*)] [*if*] [*in*] [, `rts`(*string*) `base`(*string*) `ref`(*varname*) `tename`(*newvarname*) `noprint`]

Specification

outputs is the list of output variables.

inputs is the list of input variables.

3.2 Options for `tenonradial`

Technology

`rts`(*rtsassumption)* specifies returns to scale assumption.

Specifying `rts`(*crs*) requests that measure of technical efficiency is computed under the assumption of constant returns to scale. `rts`(*crs*) is the default.

Specifying `rts`(*nrs*) requests that measure of technical efficiency is computed under the assumption of non-increasing returns to scale.

Specifying `rts`(*vrs*) requests that measure of technical efficiency is computed under the assumption of variable returns to scale.

`base`(*basetype)* specifies type of optimization.

Specifying `base`(*output*) requests that output-based measure is computed. `base`(*output*) is the default.

Specifying `base`(*input*) requests that input-based measure is computed.

Reference Set

ref_outputs is the optional list of output variables for the reference set.

The number of variables in *ref_outputs* must be equal to the number of variables in *outputs*.

If *ref_outputs* is specified *ref_inputs* must also be specified.

If *ref_outputs* is not specified *ref_inputs* must also not be specified.

ref_outputs = *ref_inputs* must be enclosed in parentheses: (*ref_outputs* = *ref_inputs*).

ref_inputs is the optional list of input variables for the reference set.

The number of variables in *ref_inputs* must be equal to the number of variables in *inputs*.

If *ref_inputs* is specified *ref_outputs* must also be specified.

If *ref_inputs* is not specified *ref_outputs* must also not be specified.

ref_outputs = *ref_inputs* must be enclosed in parentheses: (*ref_outputs* = *ref_inputs*).

`ref(`*varname*`)` specifies the indicator variable *varname* that defines which data points of *outputs* and *inputs* form the technology reference set. If *ref_outputs* and *ref_inputs* are specified *varname* defines which data points of *ref_outputs* and *ref_inputs* form the technology reference set.

Variable generation

`tenames(`*newvarname*`)` creates *newvarname* containing the nonradial measures of technical efficiency.

Miscellaneous

`noprint` suppresses the estimation details, description of the data and reference set.

3.3 Output, generated variable, and saved results

If `noprint` is not specified, `tenonradial` produces the summary of the model, data, and note about the reference set. Specifying `tenames(`*newvarname*`)` will generate *newvarname* containing the nonradial measures of technical efficiency in the current data set.

`tenonradial` stores the following to `e()`:

Macros			
`e(title)`	title in estimation output	`e(outputs)`	the list of output variables
`e(cmd)`	`tenonradial`	`e(inputs)`	the list of input variables
`e(cmdline)`	command as typed	`e(ref_outputs)`	the list of output variables
`e(rts)`	CRS, NRS, or VRS	`e(ref_inputs)`	the list of input variables
`e(base)`	output or input		
Scalars			
`e(K)`	number of data points	`e(Kref)`	number of data points in the reference set
`e(M)`	number of outputs	`e(N)`	number of inputs
`e(reps)`	number of bootstrap replications		
Matrices			
`e(te)`	Kx1 matrix with measures of technical efficiency		
Functions			
`e(sample)`	marks estimation sample		

4 The `teradial` command

The syntax, options, output, generated variable, and saved results are identical to those of `tenonradial`.

5 The `teradialbc` command

`teradialbc` performs statistical inference about the radial measure of technical efficiency

5.1 Syntax

`teradialbc` *outputs* `=` *inputs* [`(`*ref_outputs* `=` *ref_inputs*`)`] [*if*] [*in*] [`,`
`rts(`*string*`)` `base(`*string*`)` `ref(`*varname*`)` `subsampling` `kappa(`#`)` `smoothed`
`heterogeneous` `reps(`#`)` `level(`#`)` `tename(`*newvarname*`)` `tebc(`*newvarname*`)`
`biasboot(`*newvarname*`)` `varboot(`*newvarname*`)` `biassqvar(`*newvarname*`)`
`telower(`*newvarname*`)` `teupper(`*newvarname*`)` `noprint` `nodots`]

5.2 Options for `teradialbc`

Technology

identical to those of `tenonradial`.

Reference Set

identical to those of `tenonradial`.

Bootstrap

`subsampling` requests that the reference set is bootstrapped with subsampling. If `subsampling` is not specified, bootstrap with smoothing is used.

`kappa(`#`)` sets the size of the subsample as $K^{\mathtt{kappa}}$, where K is the number of data points in the original reference set. The default value is 0.7. # may be between 0.5 and 1.

`smoothed` requests that the reference set is bootstrapped with smoothing. Option `smoothed` must not be satisfied: if `subsampling` is not specified, bootstrap with smoothing is used. This option is rather for keeping track of the bootstrap type.

`heterogeneous` requests that the reference set is bootstrapped with heterogeneous smoothing. If `heterogeneous` is not specified, homogeneous smoothed bootstrap is used.

`reps(`#`)` specifies the number of bootstrap replications to be performed. The default is 999. The minimum is 200. Adequate estimates of confidence intervals using bias-corrected methods typically require 1,000 or more replications.

Statistical inference

`level(`#`)` sets confidence level; default is `level(95)`.

Variable generation

<u>te</u>names(*newvarname*) creates *newvarname* containing the radial measures of technical efficiency.

tebc(*newvarname*) creates *newvarname* with bias-corrected radial measures of technical efficiency.

<u>bias</u>boot(*newvarname*) creates *newvarname* with bootstrap bias estimate for the original radial measures of technical efficiency.

<u>var</u>boot(*newvarname*) creates *newvarname* with bootstrap variance estimate for the radial measures of technical efficiency.

biassqvar(*newvarname*) creates *newvarname* with three times the ratio of bias squared to variance for radial measures of technical efficiency.

<u>telow</u>er(*newvarname*) creates *newvarname* with the lower bound estimate for radial measures of technical efficiency.

<u>teupp</u>er(*newvarname*) creates *newvarname* with the upper bound estimate for radial measures of technical efficiency.

Miscellaneous

noprint suppresses display of the log.

nodots suppresses display of the replication dots. One dot character is displayed for each successful replication.

5.3 Details

teradialbc performs bias correction of the radial output- or input-based measure of technical efficiency under the assupmtion of constant, non-increasing, or variable returns to scale technology, computes bias and constructs confidence intervals.

If reference set is not specified, the reference set is formed by data points, for which measures of technical efficiency are computed.

Statistical inference (computation of bias, variance, and confidence interval) is performed for data points where the real number of bootstrap replications is at least 100. Matrix e(realreps) saves real number of bootstrap replications, which may be smaller than reps(#).

If at least one input-based bias-corrected Farrell measure of technical efficiency is negative, the analysis and statistical inference is performed in terms of Shephard distance functions, a reciprocal of the Debreu-Farrell measure.

5.4 Dependency of `teradialbc`

`teradialbc` depends on mata functions `kdens_bw()` and `mm_quantile()`. If not installed, install by typing `-net install kdens.pkg-` and `-ssc install moremata-`.

5.5 Output, generated variables, and saved results

If `noprint` and `nodots` are not specified, `teradialbc` produces the summary of the model, data, note about the reference set, and dislays replication dots. Several variables related to statistical inference can be generated in the current data set. For example, specifying `tenames(`*newvarname*`)` will generate *newvarname* containing the nonradial measures of technical efficiency. Options above describe creating other variables.

`teradialbc` stores the following to `e()`:

Macros			
`e(title)`	title in estimation output	`e(outputs)`	the list of output variables
`e(cmd)`	`teradialbc`	`e(inputs)`	the list of input variables
`e(cmdline)`	command as typed	`e(ref_outputs)`	the list of output variables for the reference set
`e(rts)`	CRS, NRS, or VRS	`e(ref_inputs)`	the list of input variables for the reference set
`e(base)`	output or input		
Scalars			
`e(K)`	number of data points	`e(Kref)`	number of data points in the reference set
`e(M)`	number of outputs	`e(reps)`	number of bootstrap replications
`e(N)`	number of inputs		
Matrices			
`e(te)`	Kx1 matrix with measures of technical efficiency	`e(telow)`	Kx1 matrix with the lower bound estimate for radial measures of technical efficiency
`e(tebc)`	Kx1 matrix with bias-corrected radial measures of technical efficiency	`e(teupp)`	Kx1 matrix with the upper bound estimate for radial measures of technical efficiency
`e(biasboot)`	Kx1 matrix with bootstrap bias estimate for the original radial measures of technical efficiency	`e(biassqvar)`	Kx1 matrix with three times the ratio of bias squared to variance for radial measures of technical efficiency
`e(varboot)`	Kx1 matrix with bootstrap variance estimate for the radial measures of technical efficiency	`e(realreps)`	Kx1 matrix with number of bootstrap replications that were used for statistical inference
Functions			
`e(sample)`	marks estimation sample		

6 The `nptestind` command

`nptestind` performs nonparametric test of independence.

6.1 Syntax

`nptestind` *outputs* `=` *inputs* [*if*] [*in*] [`, rts(`*string*`) base(`*string*`) reps(`#`)`
`alpha(`#`) noprint nodots`]

6.2 Options for `nptestind`

Technology

identical to those of `tenonradial`.

Bootstrap

`reps(#)` specifies the number of bootstrap replications to be performed. The default is 999. The minimum is 200. Adequate estimates of confidence intervals using bias-corrected methods typically require 1,000 or more replications.

Statistical inference

`alpha(#)` sets significance level; default is `alpha(0.05)`.

Miscellaneous

identical to those of `teradialbc`.

6.3 Dependency of `nptestind`

`nptestind` depends on mata function `kdens_bw()`. If not installed, install by typing `-net install kdens.pkg-`.

6.4 Output and saved results

If `noprint` and `nodots` are not specified, `nptestind` produces the summary of the model, data, note about the reference set, and dislays replication dots. Several variables related to statistical inference can be generated in the current data set.

`nptestind` stores the following to `e()`:

Macros

`e(title)`	title in estimation output	`e(cmd)`	`teradialbc`
`e(cmdline)`	command as typed	`e(reps)`	number of bootstrap replications
`e(rts)`	CRS, NRS, or VRS	`e(outputs)`	the list of output variables
`e(base)`	output or input	`e(inputs)`	the list of input variables

Scalars

`e(K)`	number of data points	`e(t4n)`	T4n statistic
`e(M)`	number of outputs	`e(pvalue)`	*p*-value of the test that the measure of technical efficiency and mix of inputs (or outputs) are independent
`e(N)`	number of inputs		

Matrices

`e(t4nboot)`	reps x 1 matrix with bootstrap values of the T4n statistic

7 The `nptestrts` command

`nptestrts` performs nonparametric test of returns to scale.

7.1 Syntax

`nptestrts` *outputs* `=` *inputs* [*if*] [*in*] [`,` `rts(`*string*`)` `base(`*string*`)` `ref(`*varname*`)` `heterogeneous` `reps(`*#*`)` `alpha(`*#*`)` `testtwo` `tecrsname(`*newvarname*`)` `tenrsname(`*newvarname*`)` `tevrsname(`*newvarname*`)` `sefficiency(`*newvarname*`)` `psefficient(`*newvarname*`)` `sefficient(`*newvarname*`)` `nrsovervrs(`*newvarname*`)` `pineffdrs(`*newvarname*`)` `sineffdrs(`*newvarname*`)` `noprint` `nodots`]

7.2 Options for `nptestrts`

Technology

identical to those of `tenonradial`, except for `rts`.

Reference Set

identical to those of `tenonradial`.

Bootstrap

`heterogeneous` requests that the reference set is bootstrapped with heterogeneous smoothing. If `heterogeneous` is not specified, homogeneous smoothed bootstrap is used.

`reps(`*#*`)` specifies the number of bootstrap replications to be performed. The default is 999. The minimum is 200. Adequate estimates of confidence intervals using bias-corrected methods typically require 1,000 or more replications.

Statistical inference

`alpha(#)` sets significance level; default is `alpha(0.05)`.

`testtwo` specifies that the Test #2 is performed.

If `testtwo` is not specified, `nptestrts` performs only Test #1, which consists of two parts. First, the null hypothesis that the technology is globally CRS (vs VRS) is tested. Second, the null hypothesis that the data point is scale efficient is tested.

If `testtwo` is specified, `nptestrts` may perform Test #2. If the null hypothesis that the technology is CRS is rejected, `testtwo` requests that `nptestrts` tests the null hypothesis that the technology is NIRS (vs VRS). If not all data points are scale efficient, `nptestrts` tests that the reason for scale inefficiency is operating under decreasing returns to scale (DRS). If the null hypothesis that the technology is CRS is not rejected and all data points are scale efficient, `nptestrts` will not perform Test #2 even if `testtwo` is specified.

Variable generation

`tenames(`*newvarname*`)` creates *newvarname* containing the radial measures of technical efficiency.

`tecrsnames(`*newvarname*`)` creates *newvarname* containing the radial measures of technical efficiency under the assumption of CRS.

`tenrsnames(`*newvarname*`)` creates *newvarname* containing the radial measures of technical efficiency under the assumption of NIRS.

`tevrsnames(`*newvarname*`)` creates *newvarname* containing the radial measures of technical efficiency under the assumption of VRS.

`sefficiency(`*newvarname*`)` creates *newvarname* containing scale efficiency, the ratio of measures of technical efficiency under CRS and VRS.

`psefficient(`*newvarname*`)` creates *newvarname* containing *p*-value of the test that data point is statistically scale efficient.

`sefficient(`*newvarname*`)` creates indicator *newvarname* equal one if statistically scale efficient.

`nrsovervrs(`*newvarname*`)` creates *newvarname* containing the ratio of measures of technical efficiency under NIRS and VRS.

`pineffdrs(`*newvarname*`)` creates *newvarname* containing *p*-value of the test that data point is scale inefficient due to operating under DRS.

`sineffdrs(`*newvarname*`)` creates indicator *newvarname* equal one if statistically scale inefficient due to operating under DRS.

Miscellaneous

`noprint` suppresses display of the log.

`nodots` suppresses display of the replication dots. One dot character is displayed for each successful replication.

7.3 Details

`nptestrts` performs nonparametric test of returns to scale.

If `testtwo` is not specified, `nptestrts` performs only Test #1, which consists of two parts. First, the null hypothesis that the technology is globally CRS (vs VRS) is tested. Second, the null hypothesis that the data point is scale efficient is tested.

If `testtwo` is specified, `nptestrts` may perform Test #2. If the null hypothesis that the technology is CRS is rejected, opt testtwo requests that `nptestrts` tests the null hypothesis that the technology is NIRS (vs VRS). If not all data points are scale efficient, `nptestrts` tests that the reason for scale inefficiency is operating under DRS. If the null hypothesis that the technology is CRS is not rejected and all data points are scale efficient, `nptestrts` will not perform Test #2 even if `testtwo` is specified.

7.4 Dependency of `nptestrts`

`nptestrts` depends on mata function `kdens_bw()`. If not installed, install by typing `-net install kdens.pkg-`.

7.5 Output, generated variables, and saved results

If `noprint` and `nodots` are not specified, `nptestrts` produces the summary of the model, data, note about the reference set, and dislays replication dots. Several variables related to nonparametric test can be generated in the current data set. For example, specifying `tecrsnames(`*newvarname*`)` will generate *newvarname* containing the nonradial measures of technical efficiency under the assumption of constant returns to scale. Options above describe creating other variables.

`teradialbc` stores the following to `e()`:

Macros			
`e(title)`	title in estimation output	`e(cmd)`	`nptestrts`
`e(cmdline)`	command as typed	`e(base)`	output or input
`e(outputs)`	the list of output variables	`e(inputs)`	the list of input variables
`e(smoothtype)`	homogeneous or heterogenous		
Scalars			
`e(K)`	number of data points	`e(sefficiencyMean)`	ratio of means of technical efficiency measures under CRS and VRS
`e(M)`	number of outputs	`e(pGlobalCRS)`	p-value of the test that the technology is globally CRS
`e(N)`	number of inputs	`e(nsefficient)`	number of scale efficient data points
`e(nrsOVERvrsMean)`	ratio of means of technical efficiency measures under NIRS and VRS (if opt testtwo)	`e(pGlobalNRS)`	p-value of the test that the technology is globally NIRS (if opt testtwo)
`e(reps)`	number of bootstrap replications		
Matrices			
`e(tecrsname)`	Kx1 matrix with measures of technical efficiency under the assumption of CRS	`e(sefficiency)`	Kx1 matrix containing scale efficiency
`e(tenrsname)`	Kx1 matrix with measures of technical efficiency under the assumption of NIRS	`e(psefficient)`	Kx1 matrix containing p-value of the test that data point is statistically scale efficient
`e(tevrsname)`	Kx1 matrix with measures of technical efficiency under the assumption of VRS	`e(sefficient)`	Kx1 matrix containing ones if statistically scale efficient
`e(pineffdrs)`	Kx1 matrix containing p-value of the test that data point is scale inefficient due to DRS (if opt testtwo)	`e(sineffdrs)`	Kx1 matrix containing ones if statistically scale inefficient due to DRS (if opt testtwo)
`e(nrsovervrs)`	Kx1 matrix containing the ratio of measures of technical efficiency under NiRS and VRS (if opt testtwo)		
Functions			
`e(sample)`	marks estimation sample		

8 Empirical application

In this section, we show how to use new commands and interpret the output based on two widely used data sets.

8.1 Data: CCR81

The first dataset comes from Charnes et al. (1981). The data were originally used to evaluate the efficiency of public programs and their management. In what follows, we stick to output-based efficiency measurement.

We artificially create a variable `dref` to illustrate the capabilities of new commands. We do not suppress the estimation details, description of the data and reference set for output-based radial measure of technical efficiency under the assumption of CRS technology. We do so for the remaining of radial and all of nonradial measures. Finally, we list the measures for first 7 observations:

```
. use ccr81, clear
(Program Follow Through at 70 US Primary Schools)
. generate dref = x5 != 10
. teradial y1 y2 y3 = x1 x2 x3 x4 x5, r(c) b(o) ref(dref) tename(TErdCRSo)
 Radial (Debreu-Farrell) output-based measures of technical efficiency under assumption
 of CRS technology are computed for the following data:
    Number of data points (K) = 70
    Number of outputs     (M) = 3
    Number of inputs      (N) = 5
 Reference set is formed by 68 provided reference data points
. teradial y1 y2 y3 = x1 x2 x3 x4 x5, r(n) b(o) ref(dref) tename(TErdNRSo) nopr
> int
. teradial y1 y2 y3 = x1 x2 x3 x4 x5, r(v) b(o) ref(dref) tename(TErdVRSo) nopr
> int
. tenonradial y1 y2 y3 = x1 x2 x3 x4 x5, r(c) b(o) ref(dref) tename(TEnrCRSo) n
> oprint
. tenonradial y1 y2 y3 = x1 x2 x3 x4 x5, r(n) b(o) ref(dref) tename(TEnrNRSo) n
> oprint
. tenonradial y1 y2 y3 = x1 x2 x3 x4 x5, r(v) b(o) ref(dref) tename(TEnrVRSo) n
> oprint
. list TErdCRSo TErdNRSo TErdVRSo TEnrCRSo TEnrNRSo TEnrVRSo in 1/7
```

	TErdCRSo	TErdNRSo	TErdVRSo	TEnrCRSo	TEnrNRSo	TEnrVRSo
1.	1.087257	1.032294	1.032294	1.11721	1.05654	1.05654
2.	1.110133	1.109314	1.109314	1.383089	1.277123	1.277123
3.	1.079034	1.068429	1.068429	1.17053	1.116582	1.116582
4.	1.119434	1.107413	1.107413	1.489086	1.471301	1.471301
5.	1.075864	1.075864	1	1.196779	1.196779	1
6.	1.107752	1.107752	1.105075	1.380214	1.378378	1.378378
7.	1.125782	1.119087	1.119087	1.575288	1.547186	1.547186

`teradial` and `tenonradial` compute measures of technical efficiency for all 70 data points using the reference set based on the restriction `x5 != 10`, which leaves two data points out. As expected, the radial measures are at least not worse than nonradial measures for each of returns to scale assumption. Figure 3 visualizes this observation. This indicates that there are slacks in outputs. Besides, for radial and nonradial measures, the measures under VRS are at least not worse than those under NIRS. The measures under NIRS are at least not worse that those under CRS. For data points 1, 2, 3, 4, and 7, measures under NIRS and VRS are equal. For data points 5 and 6, measures under NIRS and CRS are equal. We come back to scale analysis shortly when we discuss `nptestrts` command.

Before `teradialbc` is run, we need to know what type of bootstrap to employ. We perform therefore the nonparametric test of independence by running the new command `nptestind`. For illustration purposes, we run the test for all returns to scale assumption for both output- and input-based frontier models. We show the output only for the

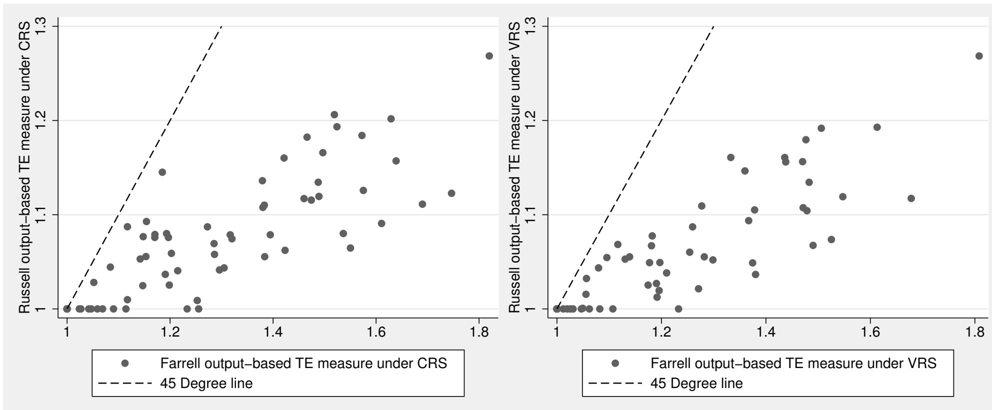

Figure 3: Scatterplot of Debreu-Farrell and Russell measures of technical efficiency under the assumption of CRS (left panel) and under the assumption of VRS (right panel)

output-based model under the assumption of CRS technology. We suppress the log for the remaining five models:

```
. matrix testsindpv = J(2, 3, .)
. matrix colnames testsindpv = CRS NiRS VRS
. matrix rownames testsindpv = output-based input-based
. nptestind y1 y2 y3 = x1 x2 x3 x4 x5, r(c) b(o) reps(999) a(0.05)
 Radial (Debreu-Farrell) output-based measures of technical efficiency under
 assumption of CRS, NIRS, and VRS technology are computed for the following
 data:
    Number of data points (K) = 70
    Number of outputs     (M) = 3
    Number of inputs      (N) = 5
 Reference set is formed by 70 data points, for which measures of technical
 efficiency are computed.
 ---------------------------------------------------------------------------
         Test
 Ho: T4n = 0 (radial (Debreu-Farrell) output-based measure of technical
 efficiency under assumption of CRS technology and mix of outputs are
 independent)
 Bootstrapping test statistic T4n (999 replications)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
(dots omitted)
..................................................   950
.................................................
 p-value of the Ho that T4n = 0 (Ho that radial (Debreu-Farrell) output-based
 measure of technical efficiency under assumption of CRS technology and mix of
 outputs are independent) = 0.0621:
 hat{T4n} = 0.0310 is not statistically greater than 0 at the 5% significance
 level
```

```
 Heterogeneous bootstrap should be used when performing output-based technical
 efficiency measurement under assumption of CRS technology
. matrix testsindpv[1,1] = e(pvalue)
. nptestind y1 y2 y3 = x1 x2 x3 x4 x5, r(n) b(o) reps(999) a(0.05) noprint
. matrix testsindpv[1,2] = e(pvalue)
. nptestind y1 y2 y3 = x1 x2 x3 x4 x5, r(v) b(o) reps(999) a(0.05) noprint
. matrix testsindpv[1,3] = e(pvalue)
. nptestind y1 y2 y3 = x1 x2 x3 x4 x5, r(c) b(i) reps(999) a(0.05) noprint
. matrix testsindpv[2,1] = e(pvalue)
. nptestind y1 y2 y3 = x1 x2 x3 x4 x5, r(n) b(i) reps(999) a(0.05) noprint
. matrix testsindpv[2,2] = e(pvalue)
. nptestind y1 y2 y3 = x1 x2 x3 x4 x5, r(v) b(i) reps(999) a(0.05) noprint
. matrix testsindpv[2,3] = e(pvalue)
. matrix list testsindpv
testsindpv[2,3]
                    CRS        NiRS         VRS
output-based  .06206206  .21921922  .03803804
 input-based  .02402402  .00500501  .24624625
```

Depending on the assumption about the technology and base of measurement, the `nptestind` concludes differently about the type of the bootstrap. In output-based efficiency measurement, the independence assumption is rejected at the 5% significance level only for VRS technology. In input-based efficiency measurement, it is only for VRS technology that the independence assumption is not rejected at the 5% significance level.

`teradialbc` can provide statistical inference for three types of bootstrap: (i) smoothed homogeneous, (ii) smoothed heterogenous, and (iii) subsampling [heterogeneous] bootstrap.

We performed each of these types under the assumption of VRS technology. The results of `nptestind` indicate that the heterogeneous bootstrap should be used, so the results for the homogeneous bootstrap cannot be trusted. We report them here for illustration purposes only. By using options `tebc`, `biassqvar`, `telower`, and `teupper`, we generate new variables in the current data set that contain bias-corrected output-based measure of technical efficiency, the statistic that compares bias and variance of the bootstrap (we also report its summary right after the command), and lower and upper bounds of 95% confidence interval for each of three types of bootstrap. Table 1 lists the measures for the first 34 data points. We let `teradialbc` output the log and bootstrap dots for the first type, but suppress them for other two:

```
. teradialbc y1 y2 y3 = x1 x2 x3 x4 x5, r(v) b(o) ref(dref) reps(999) tebc(TErd
> VRSoBC1) biassqvar(TErdVRSoBC1bv) telower(TErdVRSoLB1) teupper(TErdVRSoUB1)
 Radial (Debreu-Farrell) output-based measures of technical efficiency under
 assumption of VRS technology are computed for the following data:
    Number of data points (K) = 70
    Number of outputs     (M) = 3
    Number of inputs      (N) = 5
 Reference set is formed by 68 provided reference data points.
```

```
Bootstrapping reference set formed by 68 provided reference data points and
computing radial (Debreu-Farrell) output-based measures of technical efficiency
under assumption of VRS technology for each of 70 data points relative to the
bootstrapped reference set
Smoothed homogeneous bootstrap (999 replications)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
(dots omitted)
.................................................
. su TErdVRSoBC1bv

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
TErdVRSoBC~v |         70    3.591606    1.242244   1.922622   7.885657

. teradialbc y1 y2 y3 = x1 x2 x3 x4 x5, het r(v) b(o) ref(dref) reps(999) tebc(
> TErdVRSoBC2) biassqvar(TErdVRSoBC2bv) telower(TErdVRSoLB2) teupper(TErdVRSoUB
> 2) noprint
. su TErdVRSoBC2bv

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
TErdVRSo~2bv |         58    40.84578    47.09781   5.852742   294.3301

. teradialbc y1 y2 y3 = x1 x2 x3 x4 x5, subs r(v) b(o) ref(dref) reps(999) tebc
> (TErdVRSoBC3) biassqvar(TErdVRSoBC3bv) telower(TErdVRSoLB3) teupper(TErdVRSoU
> B3) noprint
. su TErdVRSoBC3bv

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
TErdVRSo~3bv |         70    3.644745    2.421754          0   15.34533
```

Statistic BV in Table 1 is three times the ratio of bias squared to variance of the bootstrap values of the radial measures of technical efficiency. Bias correction and statistical inference should be performed only if this statistic is well above unity. For all three types of bootstrap, BV measure is satisfactory. For smoothed homogeneous and subsampling bootstrap the values of BV are much smaller that those for smoothed heterogeneous bootstrap. If BV is small, the variance of the bootstrap values is relatively large and the mean-square error of the bias-corrected estimate of technical efficiency measure is much higher than that of the original measure. We also know from running `nptestind` that the results from smoothed homogeneous bootstrap should not be trusted.

Panels 'Smoothed heterogeneous' and 'Subsampling' in table 1 show the results for heterogeneous bootstrap. Statistic BV indicates that subsampling bootstrap introduces quite some noise. Bias-corrected measure is estimated very imprecise and results for subsampling bootstrap should not be used. This leaves us with reliable statistical inference using heterogeneous smoothed bootstrap. Here the BV statistic is well above unity.

The bias for heterogeneous smoothed bootstrap is larger than that of homogeneous smoothed bootstrap, which is implied by larger bias-corrected estimates of efficiency

Table 1: Statistical inference about the radial output-based measure of technical efficiency under the assumption of VRS

#	TE[a]	Smoothed homogeneous				Smoothed heterogeneous				Subsampling			
		BC[b]	BV[c]	LB[d]	UB[e]	BC	BV	LB	UB	BC	BV	LB	UB
1	1.032	1.056	0.290	1.033	1.129	1.177	7.362	1.124	1.290	1.153	0.395	1.040	1.414
2	1.109	1.126	0.438	1.110	1.162	1.222	3.018	1.154	1.339	1.186	0.418	1.114	1.334
3	1.068	1.087	0.310	1.069	1.141	1.158	5.647	1.115	1.220	1.138	0.642	1.068	1.250
4	1.107	1.116	0.516	1.108	1.136	1.142	2.301	1.123	1.179	1.135	0.671	1.109	1.171
5	1.000	1.048	0.275	1.000	1.194	.	.	.	.	1.000	1.700	1.000	1.000
6	1.105	1.123	0.441	1.105	1.161	1.324	2.176	1.200	1.804	1.195	0.160	1.117	1.836
7	1.119	1.129	0.555	1.120	1.146	1.154	2.526	1.134	1.187	1.149	0.749	1.122	1.180
8	1.104	1.126	0.244	1.105	1.207	1.361	5.364	1.251	1.673	1.268	0.362	1.107	1.656
9	1.161	1.174	0.476	1.161	1.200	1.201	2.148	1.175	1.241	1.216	0.788	1.163	1.274
10	1.055	1.076	0.299	1.055	1.137	1.173	6.741	1.128	1.258	1.147	0.425	1.057	1.346
11	1.000	1.034	0.396	1.000	1.110	1.137	5.790	1.088	1.262	1.080	0.295	1.000	1.284
12	1.000	1.032	0.368	1.001	1.101	1.166	4.394	1.093	1.346	1.091	0.275	1.000	1.332
13	1.156	1.164	0.729	1.156	1.177	1.183	1.524	1.165	1.218	1.192	0.389	1.159	1.250
14	1.016	1.033	0.361	1.016	1.078	.	.	.	.	1.047	0.159	1.018	1.156
15	1.000	1.050	0.302	1.000	1.195	.	.	.	.	1.175	0.274	1.000	1.525
16	1.052	1.070	0.238	1.052	1.129	1.244	3.587	1.123	1.469	1.188	0.270	1.052	1.492
17	1.000	1.048	0.275	1.001	1.194	.	.	.	.	1.222	0.337	1.000	1.657
18	1.000	1.032	0.504	1.000	1.087	1.107	2.954	1.050	1.235	1.065	0.306	1.000	1.218
19	1.049	1.066	0.328	1.050	1.111	1.122	2.470	1.076	1.198	1.122	0.296	1.050	1.305
20	1.000	1.047	0.333	1.001	1.164	1.288	6.666	1.205	1.705	1.136	0.323	1.000	1.478
21	1.000	1.042	0.377	1.001	1.128	1.170	5.913	1.108	1.336	1.111	0.368	1.000	1.350
22	1.000	1.020	0.613	1.000	1.051	1.204	2.034	1.091	1.680	1.061	0.146	1.000	1.301
23	1.025	1.036	0.285	1.026	1.069	1.069	1.578	1.035	1.121	1.098	0.243	1.027	1.259
24	1.000	1.048	0.360	1.000	1.159	1.354	4.846	1.224	2.241	1.136	0.275	1.000	1.952
25	1.021	1.030	0.938	1.022	1.041	1.047	1.697	1.031	1.075	1.044	0.472	1.028	1.082
26	1.060	1.069	0.521	1.061	1.088	1.095	1.599	1.070	1.137	1.103	0.347	1.062	1.190
27	1.000	1.035	0.384	1.000	1.111	1.207	5.644	1.141	1.469	1.094	0.221	1.000	1.372
28	1.012	1.029	0.387	1.013	1.071	1.214	1.948	1.105	1.706	1.083	0.137	1.022	1.839
29	1.180	1.199	0.316	1.180	1.254	.	.	.	.	1.272	0.768	1.184	1.359
30	1.117	1.129	0.469	1.118	1.153	1.251	2.081	1.171	1.455	1.179	0.178	1.122	1.458
31	1.193	1.204	0.618	1.193	1.224	1.245	1.651	1.211	1.317	1.225	0.748	1.199	1.266
32	1.000	1.050	0.278	1.000	1.196	.	.	.	.	1.000	0.100	1.000	1.000
33	1.049	1.069	0.713	1.051	1.103	1.097	1.376	1.062	1.158	1.142	0.350	1.051	1.323
34	1.161	1.172	0.351	1.161	1.203	1.206	1.905	1.173	1.253	1.233	0.375	1.163	1.357

[a] Original output-based measures of technical efficiency under assumption of VRS technology; [b] bias-corrected radial measures of technical efficiency; [c] three times the ratio of bias squared to variance for radial measures of technical efficiency; [d] lower bound estimate for radial measures of technical efficiency; [e] upper bound estimate for radial measures of technical efficiency.

measures. This means that homogeneous bootstrap provides optimistic estimates of the bootstrapped frontier. The 95% confidence interval is also wider, but not as wide as that for subsampling bootstrap. This might be a result of large variance of the bootstrap values for subsampling bootstrap. Statistical inference cannot be provided for observations 5, 14, 15, 17, 29, and 32. The reason for this is too few bootstrap replications, where these observations lie within the bootstrap frontier, making the solution of the linear programming problem infeasible. Indeed `e(realreps)` after `teradialbc` for mentioned data points is $0(5), 1(14), 3(15), 0(17), 3(29)$, and $0(32)$.

Finally, we turn to our discussion to the new command `nptestrts` that performs nonparametric test of returns to scale and analysis of scale efficiency. We have already determined that heterogeneous smoothed bootstrap should be used for this data set. We still provide the results using the homogeneous bootstrap and emphasize the caveat of using incorrect bootstrap procedure.

```
. nptestrts y1 y2 y3 = x1 x2 x3 x4 x5, testtwo b(o) reps(999) a(0.05) se(SE_o)
> sefficient(SEffnt_hom) sineffdrs(SiDRS_hom)
 Radial (Debreu-Farrell) output-based measures of technical efficiency under
 assumption of CRS, NIRS, and VRS technology are computed for the following
 data:
    Number of data points (K) = 70
    Number of outputs     (M) = 3
    Number of inputs      (N) = 5
 Reference set is formed by 70 data points, for which measures of technical
 efficiency are computed.
 ------------------------------------------------------------------------------
         Test #1
 Ho: mean(F_i^CRS)/mean(F_i^VRS) = 1
   and
 Ho: F_i^CRS/F_i^VRS = 1 for each of 70 data point(s)
 Bootstrapping reference set formed by 70 data points and computing radial
 (Debreu-Farrell) output-based measures of technical efficiency under
 assumption of CRS and VRS technology for each of 70 data points relative to
 the bootstrapped reference set
 Smoothed homogeneous bootstrap (999 replications)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
(dots omitted)
.................................................
 p-value of the Ho that mean(F_i^CRS)/mean(F_i^VRS) = 1 (Ho that the global
 technology is CRS) = 0.0040:
 mean(hat{F_i^CRS})/mean(hat{F_i^VRS}) = 1.0164 is statistically greater than 1
 at the 5% significance level
 All data points are scale efficient
 ------------------------------------------------------------------------------
         Test #2
 Ho: mean(F_i^NiRS)/mean(F_i^VRS) = 1
 Bootstrapping reference set formed by 70 data points and computing radial
 (Debreu-Farrell) output-based measures of technical efficiency under
 assumption of CRS and VRS technology for each of 70 data points relative to
 the bootstrapped reference set
 Smoothed homogeneous bootstrap (999 replications)
```

```
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
(dots omitted)
................................................
 p-value of the Ho that mean(F_i^NiRS)/mean(F_i^VRS) = 1 (Ho that the global
 technology is NiRS) = 0.0010:
 mean(hat{F_i^NiRS})/mean(hat{F_i^VRS}) = 1.0085 is statistically greater than
 1 at the 5% significance level

. table SEffnt_hom
```

Indicator variable if statistically scale efficient	Freq.
scale efficient	70

The p-value of the null hypothesis that global technology is constant returns to scale (Test #1) using homogeneous smoothed bootstrap is very small implying CRS is not an appropriate assumption. Further, the the null hypothesis that global technology is nonincreasing returns to scale (Test #2) is also rejected. Hence, nonparametric test of returns to scale advises performing efficiency measurement under assumption of VRS technology. Additionally, the message `All data points are scale efficient` implies that the Test #1 is not rejected for a single data point. That the global returns to scale is not CRS is at odds with the latter finding. We now perform the test using heterogeneous smoothed bootstrap:

```
. nptestrts y1 y2 y3 = x1 x2 x3 x4 x5, testtwo het b(o) reps(999) a(0.05) seffi
> cient(SEffnt_het) sineffdrs(SiDRS_het)
 Radial (Debreu-Farrell) output-based measures of technical efficiency under
 assumption of CRS, NIRS, and VRS technology are computed for the following
 data:
    Number of data points (K) = 70
    Number of outputs     (M) = 3
    Number of inputs      (N) = 5
 Reference set is formed by 70 data points, for which measures of technical
 efficiency are computed.

          Test #1
 Ho: mean(F_i^CRS)/mean(F_i^VRS) = 1
   and
 Ho: F_i^CRS/F_i^VRS = 1 for each of 70 data point(s)
 Bootstrapping reference set formed by 70 data points and computing radial
 (Debreu-Farrell) output-based measures of technical efficiency under
 assumption of CRS and VRS technology for each of 70 data points relative to
 the bootstrapped reference set
 Smoothed heterogeneous bootstrap (999 replications)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
```

```
(dots omitted)
 p-value of the Ho that mean(F_i^CRS)/mean(F_i^VRS) = 1 (Ho that the global
 technology is CRS) = 1.0000:
 mean(hat{F_i^CRS})/mean(hat{F_i^VRS}) = 1.0164 is not statistically greater
 than 1 at the 5% significance level
 ______________________________________________________________________________

         Test #2
 Ho: F_i^NiRS/F_i^VRS = 1 for each of 1 scale inefficient data point(s)
 Bootstrapping reference set formed by 70 data points and computing radial
 (Debreu-Farrell) output-based measures of technical efficiency under
 assumption of NIRS and VRS technology for each of 70 data points relative to
 the bootstrapped reference set
 Smoothed heterogeneous bootstrap (999 replications)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
(dots omitted)
. table SEffnt_het
```

Indicator variable if statistically scale efficient	Freq.
scale inefficient	1
scale efficient	69

```
. table SiDRS_het
```

Indicator variable if statistically scale inefficient due to DRS	Freq.
scale inefficient due to DRS	1

Using heterogeneous smoothed bootstrap, nonparametric test fails to reject the null hypothesis that the global technology is CRS. This means there is no need to test that global technology is nonincreasing returns to scale (Test #2). Performing Test $\#1_k$, $k = 1, \ldots 70$ however suggests that one of 70 data points is scale inefficient. Since option `testtwo` was specified, Test $\#2_k$ is performed for this single data point to determine the nature of its scale inefficiency. Using option `sineffdrs` we generated an indicator variable `SiDRS_het` equal to 1 if statistically scale inefficient due to DRS. `table SiDRS_het` identifies that the data point is scale inefficient due to operating under DRS portion of the technology (such as a data point (x_j, y_j) in terms of figure 1). Table 2 lists the original measures of technical efficiency under assumption of CRS and VRS technology, scale efficiency measure, as well as indicator variables if statistically scale efficient and the nature of scale inefficiency. Consider data point 1. That it is statistically scale efficient means that 1.053 is not statistically larger than 1. That data point 5 is not statically scale efficient using heterogeneous bootstrap means that 1.076 is statistically larger than 1.

Table 2: Scale analysis

#	CRS[a]	VRS[a]	SE	Scale efficient[c] (homogeneous)	Scale efficient (heterogeneous)	Scale inefficient due to DRS[d] (heterogeneous)
1	1.087	1.032	1.053	scale efficient	scale efficient	.
2	1.110	1.109	1.001	scale efficient	scale efficient	.
3	1.079	1.068	1.010	scale efficient	scale efficient	.
4	1.119	1.107	1.011	scale efficient	scale efficient	.
5	1.076	1.000	1.076	scale efficient	scale inefficient	scale inefficient due to DRS
6	1.108	1.105	1.002	scale efficient	scale efficient	.
7	1.126	1.119	1.006	scale efficient	scale efficient	.
8	1.111	1.104	1.006	scale efficient	scale efficient	.
9	1.184	1.161	1.020	scale efficient	scale efficient	.
10	1.077	1.055	1.021	scale efficient	scale efficient	.
11	1.025	1.000	1.025	scale efficient	scale efficient	.
12	1.028	1.000	1.028	scale efficient	scale efficient	.
13	1.166	1.156	1.009	scale efficient	scale efficient	.
14	1.076	1.016	1.059	scale efficient	scale efficient	.
15	1.000	1.000	1.000	scale efficient	scale efficient	.
16	1.065	1.052	1.012	scale efficient	scale efficient	.
17	1.000	1.000	1.000	scale efficient	scale efficient	.
18	1.000	1.000	1.000	scale efficient	scale efficient	.
19	1.058	1.049	1.008	scale efficient	scale efficient	.
20	1.000	1.000	1.000	scale efficient	scale efficient	.
21	1.000	1.000	1.000	scale efficient	scale efficient	.
22	1.000	1.000	1.000	scale efficient	scale efficient	.
23	1.044	1.025	1.018	scale efficient	scale efficient	.
24	1.000	1.000	1.000	scale efficient	scale efficient	.
25	1.041	1.021	1.020	scale efficient	scale efficient	.
26	1.074	1.060	1.013	scale efficient	scale efficient	.
27	1.000	1.000	1.000	scale efficient	scale efficient	.
28	1.059	1.012	1.046	scale efficient	scale efficient	.
29	1.206	1.180	1.023	scale efficient	scale efficient	.
30	1.123	1.117	1.005	scale efficient	scale efficient	.
31	1.202	1.193	1.007	scale efficient	scale efficient	.
32	1.117	1.000	1.117	scale efficient	scale efficient	.
33	1.079	1.049	1.028	scale efficient	scale efficient	.
34	1.182	1.161	1.019	scale efficient	scale efficient	.

[a] Measure of technical efficiency under the assumption of CRS;
[b] Measure of technical efficiency under the assumption of VRS; [c] Statistically scale efficient;
[d] Statistically scale inefficient due to DRS.

The nonparametric test of returns to scale concludes differently about the null hypothesis depending on type of the bootstrap. This should not be surprising as it is a consequence of employing an inconsistent bootstrap procedure. The results of the test are based on correct mimicking of the data generating process. If this is not guaranteed, the bootstrap procedure is inconsistent and results of the nonparametric test cannot be trusted. For this particular data and base of efficiency measurement, smoothed heterogeneous bootstrap should be employed.

8.2 Data: PWT5.6

The second dataset, the Penn World Tables were used by Kumar and Russell (2002) among others. See Heston and Summers (1991) for more details on the data set. The purpose of this short study is to construct Malmquist productivity index (MPI) between 1965 and 1990, and perform analysis of productivity change by decomposing the MPI. Malmquist Productivity Index makes use of the output distance function which is the reciprocal of the Debreu-Farrell measure of technical efficiency (Caves et al. 1982; Färe et al. 1994a).

The Malmquist output-based productivity index MPI from time-period b to time period c for data point k is given by:

$$MPI_k^{o,bc} = \left[\frac{F^o(y_{k,b}, x_{k,b}, y_b, x_b|\text{CRS})}{F^o(y_{k,c}, x_{k,c}, y_b, x_b|\text{CRS})} \times \frac{F^o(y_{k,b}, x_{k,b}, y_c, x_c|\text{CRS})}{F^o(y_{k,c}, x_{k,c}, y_c, x_c|\text{CRS})} \right]^{1/2}.$$

This index maybe decomposed as

$$\begin{aligned} MPI_k^{o,bc} = {} & \frac{F^o(y_{k,b}, x_{k,b}, y_b, x_b|\text{CRS})}{F^o(y_{k,c}, x_{k,c}, y_c, x_c|\text{CRS})} \\ & \times \left[\frac{F^o(y_{k,c}, x_{k,c}, y_c, x_c|\text{CRS})}{F^o(y_{k,c}, x_{k,c}, y_b, x_b|\text{CRS})} \frac{F^o(y_{k,b}, x_{k,b}, y_c, x_c|\text{CRS})}{F^o(y_{k,b}, x_{k,b}, y_c, x_c|\text{CRS})} \right]^{1/2}, \end{aligned} \tag{17}$$

where $F^o(y_{k,d}, x_{k,d}, y_a, x_a|\text{CRS})$ is the Debreu-Farrell measure calculated for data point k in time period d to the frontier formed by observations (y_a, x_a) under the assumption of CRS technology. The first term in (17) measures the contribution of technical efficiency change to productivity change. The second term in (17) measures the contribution of technical change to productivity change:

$$MPI = EFF \times TECH.$$

If $EFF > 1$ (< 1 in input-based measurement), change in efficiency has positively contributed to productivity change from time-period b to time period c. The meaning of $TECH$ is the following: $TECH > / = / < 1$ implies that technical progress/stagnation/regress has occurred between periods b and c.

The decomposition of the Malmquist productivity index in (17) can be extended. Calculating the Debreu-Farrell measure under VRS, Malmquist productivity index can be decomposed into three components attributable to (i) Pure Technical Efficiency

Change ($PEFF$), (ii) Technological Change ($TECH$) and (iii) Scale Efficiency Change (SEC) (Färe et al. 1994b). The decomposition of the output-based MPI from time-period b to time period c for data point k is given by:

$$\begin{aligned} MPI_k^{o,bc} = & \frac{F^o(y_{k,b}, x_{k,b}, y_b, x_b|\text{VRS})}{F^o(y_{k,c}, x_{k,c}, y_c, x_c|\text{VRS})} \\ & \times \left[\frac{F^o(y_{k,c}, x_{k,c}, y_c, x_c|\text{CRS})}{F^o(y_{k,c}, x_{k,c}, y_b, x_b|\text{CRS})} \frac{F^o(y_{k,b}, x_{k,b}, y_c, x_c|\text{CRS})}{F^o(y_{k,b}, x_{k,b}, y_c, x_c|\text{CRS})} \right]^{1/2} \\ & \times \frac{S_k^o(y_{k,b}, x_{k,b})}{S_k^o(y_{k,c}, x_{k,c})}, \end{aligned} \tag{18}$$

Where S_k^o is scale efficiency defined in (11).

From the outset, it is not clear which of the decompositions, (17) or (18) should be used. We first perform the nonparametric test of returns to scale using heterogeneous bootstrap:

```
. use pwt56, clear

. reshape wide y k l, i(nu country) j(year)
(note: j = 1965 1990)

Data                               long   ->   wide
-----------------------------------------------------------------------------

Number of obs.                      104   ->      52
Number of variables                   6   ->       8
j variable (2 values)              year   ->   (dropped)
xij variables:
                                      y   ->   y1965 y1990
                                      k   ->   k1965 k1990
                                      l   ->   l1965 l1990
-----------------------------------------------------------------------------

. nptestrts y1965 = k1965 l1965, het b(o) reps(999) a(0.05)
 Radial (Debreu-Farrell) output-based measures of technical efficiency under
 assumption of CRS, NIRS, and VRS technology are computed for the following
 data:

    Number of data points (K) = 52
    Number of outputs     (M) = 1
    Number of inputs      (N) = 2

 Reference set is formed by 52 data points, for which measures of technical
 efficiency are computed.
 ------------------------------------------------------------------------------

         Test #1

 Ho: mean(F_i^CRS)/mean(F_i^VRS) = 1
   and
 Ho: F_i^CRS/F_i^VRS = 1 for each of 52 data point(s)

 Bootstrapping reference set formed by 52 data points and computing radial
 (Debreu-Farrell) output-based measures of technical efficiency under
 assumption of CRS and VRS technology for each of 52 data points relative to
 the bootstrapped reference set

 Smoothed heterogeneous bootstrap (999 replications)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
(dots omitted)
```

```
..............................................
p-value of the Ho that mean(F_i^CRS)/mean(F_i^VRS) = 1 (Ho that the global
technology is CRS) = 0.9920:
mean(hat{F_i^CRS})/mean(hat{F_i^VRS}) = 1.1196 is not statistically greater
than 1 at the 5% significance level
All data points are scale efficient
```

The null hypothesis that the global technology is CRS cannot be rejected. Besides all countries are scale efficient so that third component in (18) is essentially one. We therefore proceed with decomposition (17). We calculate the required efficiency measures

```
. teradial y1965 = k1965 l1965 (y1965 = k1965 l1965), r(c) b(o) tename(F11) nop
> rint
. teradial y1990 = k1990 l1990 (y1965 = k1965 l1965), r(c) b(o) tename(F21) nop
> rint
. teradial y1965 = k1965 l1965 (y1990 = k1990 l1990), r(c) b(o) tename(F12) nop
> rint
. teradial y1990 = k1990 l1990 (y1990 = k1990 l1990), r(c) b(o) tename(F22) nop
> rint
. g mpi = sqrt(F12 / F22 * F22 / F21)
. g effch = F11 / F22
. g techch = mpi / effch
```

and present the results of the decomposition for the first 34 out of 52 data points in Table 3. We discuss selected results.

Argentina was on the frontier in 1965 but moved away from the 1990 frontier. Hong Kong was quite inefficient in 1965 but in 1990 it defines the frontier. We also observe that productivity of industrialized countries such as Australia, Austria, and Belgium has increased, while productivity has fallen for Argentina, Bolivia, Equador among others. The productivity has increased for example in Australia due to both improved efficiency and technology. In Bolivia, the main reason for decreased productivity was loss in efficiency. In Malawi, on the contrary efficiency change has positively contributed to growth of productivity, but technology has deteriorated so much that the entire productivity has decreased.

9 Sample restriction, discussion and runtime

❏ Technical note

All functions create Stata matrices and feed them to plugin. The number of data points that can be used in all functions is thus limited by [R] **matsize**. Stata/IC allows maximum of 800, while Stata/MP and Stata/SE allow 11000 data points.

Table 3: Measures of technical efficiency and Malmquist Productivity Index

#	Country	1965[a]	1990[b]	MPI	EFFch	TECHch
1	Argentina	1.000	1.546	0.818	0.647	1.264
2	Australia	1.320	1.213	1.184	1.088	1.088
3	Austria	1.174	1.374	1.067	0.854	1.249
4	Belgium	1.419	1.159	1.247	1.225	1.018
5	Bolivia	2.002	2.457	0.948	0.815	1.163
6	Canada	1.261	1.070	1.207	1.179	1.023
7	Chile	1.180	1.549	0.889	0.762	1.167
8	Columbia	2.415	2.243	1.062	1.077	0.987
9	Denmark	1.324	1.432	1.084	0.924	1.173
10	Dominican Rep.	1.383	1.953	0.914	0.708	1.291
11	Equador	2.664	2.756	0.961	0.966	0.994
12	Finland	1.963	1.344	1.335	1.460	0.915
13	France	1.257	1.211	1.187	1.038	1.144
14	Germany, West	1.450	1.222	1.230	1.187	1.036
15	Greece	1.828	1.673	1.079	1.093	0.988
16	Guatemala	1.228	1.369	1.037	0.897	1.156
17	Honduras	2.224	2.431	1.022	0.915	1.117
18	Hong Kong	2.202	1.000	1.519	2.202	0.690
19	Iceland	1.041	1.146	0.971	0.909	1.068
20	India	2.723	2.417	1.226	1.127	1.088
21	Ireland	1.411	1.184	1.106	1.192	0.928
22	Israel	1.664	1.192	1.209	1.396	0.866
23	Italy	1.490	1.131	1.301	1.318	0.988
24	Jamaica	1.774	1.930	1.017	0.919	1.107
25	Japan	1.684	1.617	1.169	1.041	1.123
26	Kenya	3.902	3.411	1.328	1.144	1.161
27	Korea, Rep	2.309	1.632	1.225	1.415	0.866
28	Malawi	3.515	2.996	0.621	1.173	0.529
29	Mauritius	1.062	1.025	1.115	1.036	1.076
30	Mexico	1.171	1.347	0.950	0.869	1.093
31	Netherlands	1.190	1.130	1.141	1.054	1.082
32	New Zealand	1.186	1.406	1.004	0.843	1.191
33	Norway	1.628	1.257	1.492	1.295	1.152
34	Panama	2.266	3.021	0.859	0.750	1.146

[a] Measure of technical efficiency under the assumption of CRS in 1965;
[b] Measure of technical efficiency under the assumption of CRS in 1990.

❏

❏ **Technical note**

Stata 11.2 and above can be used to run all new commands. Earlier versions of Stata can probably also be used, but Stata 11.2 is the earliest version available to authors.

❏

❏ **Technical note**

Solving linear programming problems make use of *quickhull* (http://www.qhull.org/, Barber et al. (1996)) algorithm and *GLPK Version 4.55* (GNU Linear Programming Kit 2012, available at http://www.gnu.org/software/glpk/) coded in C. The required plugins are compiled from C code. Systems for which the plugins are available are MacOX, Ubuntu, and Windows.

❏

Since the linear programming is coded in low-level language, the new commands are very fast. We have recorded the time required to do the calculations in this paper. The calculations were computed on an iMac (late 2012) desktop with a 2.9 GHz processor:

```
. timer clear
. timer on 1
. use ccr81, clear
(Program Follow Through at 70 US Primary Schools)
. gen dref = x5 != 10
. tenonradial y1 y2 y3 = x1 x2 x3 x4 x5, r(c) b(o) ref(dref) tename(TErdCRSo) n
> oprint
. timer off 1
. timer on 2
. nptestind y1 y2 y3 = x1 x2 x3 x4 x5, r(c) b(o) reps(999) a(0.05) noprint
. timer off 2
. timer on 3
. teradialbc y1 y2 y3 = x1 x2 x3 x4 x5, r(v) b(o) ref(dref) reps(999) tebc(TErd
> VRSoBC1) biassqvar(TErdVRSoBC1bv) telower(TErdVRSoLB1) teupper(TErdVRSoUB1) n
> oprint
. timer off 3
. timer on 4
. teradialbc y1 y2 y3 = x1 x2 x3 x4 x5, het r(v) b(o) ref(dref) reps(999) tebc(
> TErdVRSoBC2) biassqvar(TErdVRSoBC2bv) telower(TErdVRSoLB2) teupper(TErdVRSoUB
> 2) noprint
. timer off 4
. timer on 5
. nptestrts y1 y2 y3 = x1 x2 x3 x4 x5, b(o) reps(999) a(0.05) sefficient(SEffnt
> _hom) noprint
. nptestrts y1 y2 y3 = x1 x2 x3 x4 x5, het b(o) reps(999) a(0.05) sefficient(SE
> ffnt_het) noprint
. timer off 6
```

```
. timer list
   1:      0.16 /        1 =       0.1600
   2:      2.80 /        1 =       2.8040
   3:      9.37 /        1 =       9.3670
   4:     20.23 /        1 =      20.2280
   5:     68.55 /        1 =      68.5490
   6:    782.47 /        1 =     782.4680
```

`tenonradial` is trivial and runs instantly. This remains true even if the data set is several thousands. `nptestind` is also quite fast but may slow down as sample size grows. Employing smoothed homogeneous bootstrap in `teradialbc` runs relatively quickly on small sample (9 seconds), but using heterogeneous bootstrap is more demanding (20 seconds) and the time will increase in sample size. Running nonparametric test of returns to scale, `nptestrts`, is the most involved since instead of doing calculations for each of K data points on each bootstrap replication as in `teradialbc`, the binomial test requires bootstrap replications for each of K data points independently. This is time demanding especially when smoothed heterogeneous bootstrap is used. On a sample of 70 data points, it took just above 1 minute for homogeneous and 13 minutes for heterogeneous bootstrap.

Displaying dots does not add to the output, but rather serves an indicator how long the whole bootstrap is going to take. It can be suppressed by specifying `nodots` option in each of new commands.

10 Comparison to `dea` command in Stata, Stata Journal, 10(2): 267-80)

The new command `teradial` performs radial technical efficiency analysis, which user-written command `dea` offers (Ji and Lee 2010). The latter command has two serious limitations for a practitioner. First, it is slow with even moderate data sets. We have recorded time it takes to compute input-based measure of technical efficiency under VRS using both commands for samples of 10 to 70 in steps of 10 data points:

```
. use ccr81, clear
(Program Follow Through at 70 US Primary Schools)
. rename nu dmu
. timer clear
. * number of observations: 10(10)70
. forvalues nobs = 10(10)70{
  2.  local nobs = `nobs´
  3.  local nobs2 = `nobs´ + 1
  4.  timer on `nobs´
  5.  quietly dea x1 x2 x3 x4 x5 = y1 y2 y3 in 1/`nobs´, rts(vrs) ort(in)
  6.  timer off `nobs´
  7.  timer on `nobs2´
  8.  quietly teradial y1 y2 y3 = x1 x2 x3 x4 x5 in 1/`nobs´, r(v) b(i) tename(
> TErdVRSi_`nobs´)
  9.  timer off `nobs2´
 10. }
```

```
. timer list
  10:      8.85 /        1 =       8.8470
  11:      0.00 /        1 =       0.0030
  20:     33.87 /        1 =      33.8680
  21:      0.01 /        1 =       0.0070
  30:     71.71 /        1 =      71.7090
  31:      0.02 /        1 =       0.0180
  40:    849.43 /        1 =     849.4260
  41:      0.04 /        1 =       0.0360
  50:   1067.00 /        1 =    1066.9960
  51:      0.09 /        1 =       0.0890
  60:   1839.14 /        1 =    1839.1440
  61:      0.15 /        1 =       0.1470
  70:   1990.05 /        1 =    1990.0520
  71:      0.22 /        1 =       0.2180
```

Time it takes increases for both `dea` and `teradial`, but `dea` becomes slow very quickly. For a sample size of 70 data points `dea` needs 33 minutes, while `teradial` completes in a fifth of a second. This can be a real bottleneck in actual empirical analysis. Thus, using `dea` for making statistical inference is next to infeasible.

Second, in comparison to `dea`, `teradial` can calculate the measure technical efficiency of a data point relative to the frontier defined by user by specifying `ref` option. This is required for example for analysis of productivity change demonstrated in subsection 8.2. Such analysis is not possible using `dea` command.

11 Concluding remarks

We introduce five new Stata commands that estimate and provide statistical inference in nonparametric frontier models. `tenonradial` and `teradial` calculate nonradial Russell and radial Debreu-Farrell measures of technical efficiency, respectively. The measures can be computed for different assumption about the technology, base of the analysis, as well as relative to the frontier formed by data points provided by user. These frontier models are deterministic and resulting measures are subject to sampling variation. `teradialbc` can accommodate different types of bootstrapping techniques to provide statistical inference regarding these deterministic measures. For obtaining reliable results from `teradialbc` the bootstrap type has to be chosen such that it correctly mimics the data generating process. `nptestind` provides a simple tool to determine the type of the bootstrap consistent with the data. Finally, `nptestrts` uses bootstrap to provide inference with regards to the underlying technology and performs scale analysis of each data point.

We have presented two empirical examples. In the first example, we illustrate the capabilities of new commands and discuss implications for empirical analysis. In the second example, we show that these commands can be used to analyze the changes in productivity for 52 countries form 1965 to 1990.

12 Acknowledgments

The authors are grateful to the Lebesgue Center of Mathematics who supported the work of Pavlo Mozharovskyi according to the program PIA-ANR-11-LABX-0020-01. We also thank Howard J. Newton and an anonymous reviewer for comments.

13 References

Badunenko, O., D. J. Henderson, and S. C. Kumbhakar. 2012. When, Where and How to Perform Efficiency Estimation. *Journal of the Royal Statistical Society. Series A* 175(4): 863–92.

Barber, C. B., D. P. Dobkin, and H. Huhdanpaa. 1996. The Quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software* 22(4): 469–483.

Caves, D., L. Christensen, and W. Diewert. 1982. The economic theory of index numbers and the measurement of input, output and productivity. *Econometrica* 50: 1393–1414.

Charnes, A., W. W. Cooper, and E. Rhodes. 1981. Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through. *Management Science* 27: 668–697.

Cobb, C. W., and P. H. Douglas. 1928. A Theory of Production. *American Economic Review,* Supplement 18(1): 139–165.

Coelli, T., D. P. Rao, and G. E. Battese. 2002. *An Introduction to Efficiency and Productivity Analysis*. London: Kluwer Academic Publishers.

Debreu, G. 1951. The Coefficient of Resource Utilization. *Econometrica* 19: 273–292.

Färe, R. 1988. *Fundumentals of Production Theory*. Berlin: Springer.

Färe, R., and S. Grosskopf. 1985. A nonparametric cost approach to scale efficiency. *Scandinavian Journal of Economics* 87: 594–604.

Färe, R., S. Grosskopf, and C. A. K. Lovell. 1994a. *Production Frontiers*. Cambridge, U.K.: Cambridge University Press.

Färe, R., S. Grosskopf, M. Norris, and Z. Zhang. 1994b. Technical Progress, and Efficiency Change in Industrialized Countries. *American Economic Review* 84(1): 66–83.

Färe, R., and C. A. K. Lovell. 1994. Measuring the technical efficiency of production. *Journal of Economic Theory* 19: 150–162.

Färe, R., and D. Primont. 1995. *Multi-Output Production and Duality, Theory and Applications*. Boston: Kluwer Academic Publishers.

Farrell, M. J. 1957. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)* 120(3): 253–290.

GNU Linear Programming Kit. 2012. (http://www.gnu.org/software/glpk/glpk.html).

Heston, A., and R. Summers. 1991. The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988. *The Quarterly Journal of Economics* 106: 327–368.

Ji, Y., and C. Lee. 2010. Data envelopment analysis. *Stata Journal* 10(2): 267–280. http://www.stata-journal.com/article.html?article=st0193.

Kneip, A., L. Simar, and P. Wilson. 2008. Asymptotics and Consistent Bootstraps for DEA Estimators in Nonparametric Frontier Models. *Econometric Theory* 24: 1663–1697.

Koopmans, T. C. 1951. An Analysis of Production as an Efficient Combination of Activities. In *Activity Analysis of Production and Allocation*, ed. T. C. Koopmans. New York: Wiley.

Kumar, S., and R. R. Russell. 2002. Technological Change, Technological Catch-up, and Capital Deepening: Relative Contributions to Growth and Convergence. *American Economic Review* 92(3): 527–548.

Simar, L., and P. Wilson. 1998. Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models. *Management Science* 44: 49–61.

———. 2000. A General Methodology for Bootstrapping in Nonparametric Frontier Models. *Journal of Applied Statistics* 27: 779–802.

———. 2002. Nonparametric Tests of Return to Scale. *European Journal of Operational Research* 139: 115–132.

Wilson, P. 2003. Testing Independence in Models of Productive Efficiency. *Journal of Productivity Analysis* 20: 361–390.

About the authors

Oleg Badunenko is an assistant professor at the University of Cologne. His major research interests are efficiency and productivity analysis, as well as micro- and nonparametric econometrics.

Pavlo Mozharovskyi is a postdoc researcher at Centre Henri Lebesgue, Institute of Mathematical Research of Rennes and Agrocampus Ouest in Rennes. His main research interests lie in nonparametric and computational statistics, classification, and imputation of missing data.